Psychological Trauma: Theory, Research, Practice, and Policy
2009, Vol. 1, No. 3, 173–187

1942-9681/09/$12.00 DOI: 10.1037/a0016990

Factor Structure of Posttraumatic Stress Disorder as Measured by the Impact of Event Scale–Revised: Stability Across Cultures and Time

Daniel W. King and Robert J. Orazem
Boston University and VA Boston Healthcare System

Dean Lauterbach
Eastern Michigan University

Lynda A. King
Boston University and VA Boston Healthcare System

Claire L. Hebenstreit
VA Boston Healthcare System

Arieh Y. Shalev
Hadassah University Hospital

This study examined the structure of posttraumatic stress disorder (PTSD) as measured by the Impact of Event Scale–Revised (IES-R; Weiss & Marmar, 1997), tested factorial invariance for samples of 235 Israeli emergency room patients and 306 U.S. undergraduate students, and then evaluated factorial invariance over multiple occasions within the emergency room sample. A four-factor structure representing intrusion, avoidance-numbing, hyperarousal, and sleep emerged as the preferred model. Configural invariance over groups was supported for this model. Likewise, configural invariance over occasions was demonstrated, but metric invariance was not fully supported, with variation in the loadings on the intrusion factor over time seemingly the source of misfit. Interpretations and conclusions center on sleep as a separate factor underlying the structure of the IES-R, the distinction between avoidance and numbing as a function of how the IES-R (vs. the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*) operationalizes the numbing feature of PTSD and possible shifts in the meaning of intrusion over time.

Keywords: posttraumatic stress disorder, Impact of Event Scale–Revised, confirmatory factor analysis, invariance

The *Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition, Text Revision* (*DSM–IV–TR*; American Psychiatric Association, 2000) organizes symptoms of posttraumatic stress disorder (PTSD) into *reexperiencing* (Criterion B, intrusive recollections of the trauma), *avoidance and numbing* (Criterion C, emotional withdrawal and avoidance of people or places reminiscent of the trauma), and *hyperarousal* (Criterion D, irritability, excessive vigilance, and exaggerated startle) clusters. This three-factor structure has been criticized, how-

Daniel W. King and Lynda A. King, Department of Psychology, Boston University; Department of Psychiatry, Boston University School of Medicine; National Center for Posttraumatic Stress Disorder and Massachusetts Veterans Epidemiology Research and Information Center, VA Boston Healthcare System, Boston, Massachusetts; Robert J. Orazem, National Center for Posttraumatic Stress Disorder, VA Boston Healthcare System; Department of Psychology, Boston University, Boston, Massachusetts; Claire Hebenstreit, National Center for Posttraumatic Stress Disorder and Massachusetts Veterans Epidemiology Research and Information Center, VA Boston Healthcare System, Boston, Massachusetts; Arieh Shalev Center for Traumatic Stress and the Department of Psychiatry, Hadassah University Hospital, Jerusalem, Israel.

Robert J. Orazem is now at the Minneapolis VA Medical Center. Claire L. Hebenstreit is now at the Department of Psychology, University of Denver.

This study was supported by National Institutes of Mental Health Grants R01MH68626 (Daniel W. King, Principal Investigator) and R01MH50379 (Arieh Y. Shalev, Principal Investigator).

Correspondence concerning this article should be addressed to Daniel W. King, National Center for PTSD (116B5), VA Boston Healthcare System, 150 South Huntington Ave., Boston, MA 02130. E-mail: danking@bu.edu

ever, with particular disapproval of the inclusion of avoidance and numbing symptoms within a single factor. The present study examined several conceptually and empirically supported models of PTSD as assessed by a widely used self-report measure, the Impact of Event Scale–Revised (IES-R; Weiss & Marmar, 1997). Furthermore, we considered the structural stability of PTSD across two culturally distinct samples and longitudinally across multiple time points within one of these samples. We sought to not only document the factor structure of the IES-R, but also to enhance our understanding of the PTSD construct and its symptom clusters. Knowledge of the structure of PTSD and its consistency over groups and occasions can serve a number of useful purposes: advance theory regarding the place of the disorder within the nosology of psychiatric conditions, guide thinking about diagnostic criteria and hence contribute to the development of accurate and valid assessment tools, distinguish subtleties in the presentation of the disorder across cultures and ethnic populations and over its longitudinal course, and inform the targeting of special features for prevention and treatment programs.

Research on the factor structure of PTSD initially entailed exploratory factor analysis (EFA) methodology. This literature revealed that active avoidance and emotional numbing loaded onto separate factors (Foa, Riggs, & Gershuny, 1995; King & King, 1994; Shelby, Golden-Kreutz, & Anderson, 2005; Taylor, Kuch, Koch, Crocket, & Passey, 1998). Based on these findings, King and colleagues (King, Leskin, King, & Weathers, 1998) proposed a four-factor model of PTSD that split cluster C into separate *effortful avoidance* (C1–C2) and *emotional numbing* (C3–C7) factors.[1] This model, hereafter referred to as the "four-factor numbing model," was identical to the *DSM–IV* model in all other respects and therefore maintained its *reexperiencing* (B1-B5) and *hyperarousal* (D1-D5) factors. Results from King et al.'s (1998) confirmatory factor analysis (CFA) using responses to the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995) indicated strong absolute fit to the data as well as good fit relative to several competing models. The four-factor numbing model was affirmed in subsequent CFA studies across a variety of measures and populations (Asmundson, Wright, McCreary, & Pedlar, 2003; DuHamel et al., 2004; Marshall, 2004; McWilliams, Cox, & Asmundson, 2005; Palmieri & Fitzgerald, 2005; Palmieri, Weathers, Difede, & King, 2007).

An alternative four-factor model of PTSD, referred to here as the "four-factor dysphoria model," was proposed by Simms, Watson, and Doebbeling (2002) and tested using data from the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993). This model also decomposed *DSM–IV* cluster C symptoms into distinct factors. However, it combined the emotional numbing symptoms (C3–C7) with several non-PTSD-specific cluster D symptoms (D1–D3) that are shared with other disorders (e.g., generalized anxiety disorder) to form a general distress, or *dysphoria*, factor. The remaining cluster D symptoms (D4–D5) were retained in a putatively PTSD-specific *hyperarousal* cluster. The *reexperiencing* (B1–B5) and *avoidance* (C1–C2) factors remained consistent with the four-factor numbing model. Results from Simms et al.'s (2002) CFA indicated that the four-factor dysphoria model evidenced good absolute fit and slightly better fit than the four-factor numbing model. This model was supported in subsequent CFA studies (Baschnagel, O'Connor, Colder, & Hawk, 2005; Krause, Kaltman, Goodman, & Dutton, 2007; Palmieri et al., 2007).

The relative merits of the four-factor numbing and four-factor dysphoria models remain unclear. Both have demonstrated good absolute data fit. In addition, among studies directly comparing the two models, McWilliams et al. (2005) and Palmieri and Fitzgerald (2005) garnered support for the numbing model whereas Baschnagel et al. (2005) supported the dysphoria model, albeit relying on correlated residual scores that suggest one or more additional factors. Palmieri et al. (2007) found that the numbing model attained better fit with semistructured clinical interview data and the dysphoria model achieved better fit with self-report questionnaire data and in an overall model incorporating both types of measures.

Finally, two-factor models also have received modest empirical support. Taylor and col-

[1] Throughout this article, factors are identified in italicized font using the labels that were assigned by the authors in the original work.

leagues (1998), in an EFA employing both self-report and clinician-administered measures, identified a two-factor model that included *intrusions and avoidance* (B1–B5 and C1–C2) and *numbing and hyperarousal* (C3–C7 and D1–D3) factors. The placement of symptoms D4 and D5 was ambiguous; they loaded more strongly onto the former factor in a motor vehicle accident sample and onto the latter factor in a military sample. It is interesting to note that the pattern of item loadings that emerged in the motor vehicle accident sample (referred to hereafter as the "two-factor dysphoria model") harkens to Simms et al.'s (2002) model: a general dysphoria factor and a PTSD-specific factor (reexperiencing, avoidance, and hyperarousal). This model was also supported in a study by Buckley, Blanchard, and Hickling (1998) using CAPS data, although it is noteworthy that no alternative models were tested and absolute fit was modest. On the other hand, the factor loadings that emerged in Taylor et al.'s (1998) military sample correspond with the King et al. (1998) numbing model and will be referred to here as the "two-factor numbing model."

Extant research on the factor structure of PTSD has done much to highlight key issues in the dimensionality of the construct, yet several concerns warrant further empirical attention. First, most studies have used measures (e.g., CAPS and PCL) that correspond directly to the 17 *DSM–IV* PTSD symptoms. However, the current criteria set is the result of an iterative process involving the addition, deletion, and revision of individual symptoms. It can therefore be argued that the current *DSM–IV* symptom criteria do not necessarily reflect the construct of PTSD perfectly (Joseph, 2000) and that focusing solely on indicators of these specific symptoms may stagnate the content refinement process. Measures that include unique item content while retaining content coverage of the essential features of the disorder may shed further light on the factor structure of PTSD and lead to continued diagnostic and construct refinement. The IES-R (Weiss & Marmar, 1997) is one such measure, and although it remains one of the most widely used measures of PTSD worldwide (Weiss, 2004), its factor structure remains understudied. Thus, factor analytic research on the IES-R is important for the psychometric evaluation of the instrument and for clarifying the nature of the PTSD construct itself.

Four recent CFAs of the parent Impact of Event Scale (IES; Horowitz, Wilner, & Alvarez, 1979), which contains highly similar content except its relative lack of hyperarousal items, are instructive. The IES was originally proposed to comprise two subscales, *intrusion* and *avoidance*, a structure endorsed by Shevlin, Hunt, and Robbins (2000) and van der Ploeg, Mooren, Kleber, van der Velden, and Brom (2004). Larsson (2000), in contrast, found that this two-factor model did not adequately fit the data. Accordingly, two items reflecting sleep problems and nightmares were moved from the intrusions category to a separate *sleep disturbance* factor, yielding a three-factor model (*intrusiveness*, *avoidance*, and *sleep disturbance*) that attained good model fit. Amdur and Liberzon (2001) again found that the two-factor model proved untenable, while a four-factor model that included *intrusion*, *effortful avoidance*, *emotional numbing*, and *sleep disturbance* demonstrated superior model fit. Regarding CFAs of the IES-R, Creamer, Bell, and Failla (2003) and Beck et al. (2008) proposed a three-factor (*intrusion, avoidance,* and *hyperarousal*) structure. In both cases, the three-factor model resulted in poor fit to the data. Creamer et al. proceeded with an EFA to further explore the structure of the measure, and Beck et al. relied on modification indices to adjust the fit. It is interesting to note that Beck et al.'s modification indices suggested specification of a covariance between two sleep item residuals, hinting at the existence of a separate sleep factor. Thus, prior CFAs of the original IES and the IES-R point to the possibility that items reflecting sleep problems may comprise a separate factor.

A second issue needing further research attention is the stability of measures of PTSD across cultures and language. Currently, there is little information available regarding the cross-cultural validity of existing structural models for PTSD measures. Only two studies to date have attempted to directly compare structural findings between English-speaking and non-English-speaking samples. Norris, Perilla, and Murphy (2001) found that a slightly altered version of the four-factor numbing model attained configural invariance (i.e., the same number of factors and an equivalent pattern of item loadings) across English- and Spanish-language versions of the Revised Civilian Mississippi Scale (Norris & Perilla, 1996) completed by hurricane survivors in the United States and Mexico, respectively. In addition, Marshall (2004) found that the four-factor numbing model replicated well for PCL scores in English- and Spanish-speaking subgroups within a community violence sample. Although they did not directly compare structural findings across culturally distinct groups, Sack, Seeley, and Clarke (1997) identified the four-factor numbing model as attaining good data fit in a Khmer refugee sample, indicating that the structure of PTSD in this sample resembled results from Western samples. Finally, using a sample of West and Central African refugees, Rasmussen, Smith, and Keller (2007) tested the four-factor numbing and four-factor dysphoria models against a novel four-factor model that combined intrusion with the insomnia (D1) and difficulty concentrating (D3) symptoms to form a unique *aroused intrusion* factor that may offer avenues for future inquiry. Additional research examining the cross-cultural stability of factor solutions is clearly needed (King, King, Orazem, & Palmieri, 2006; Norris & Hamblen, 2004).

A third understudied issue in the factor structure of PTSD is factorial stability over time (King et al., 2006). In an innovative study, Baschnagel et al. (2005) used the Posttraumatic Diagnostic Scale (Foa, Cashman, Jaycox, & Perry, 1997) 1 and 3 months following trauma exposure and found support for configural invariance for the four-factor dysphoria model. However, as previously noted, their reliance on correlated residuals in each of the two solutions suggests an additional factor for each occasion. In addition, metric invariance was not supported (i.e., the magnitude of factor loadings within this model differed across occasions). Using civilian PCL data, Krause et al. (2007) offered evidence for configural, metric, and phi (equal factor intercorrelations) invariance for a sample of female victims of intimate partner violence assessed while receiving services and then one year later. Their best-fitting model conformed to the four-factor dysphoria structure. Krause et al. also established invariance over two groups.

Additional PTSD factor structure research has been limited to cross-sectional designs. Invariance of factor solutions over time is necessary before differences in factor scores over occasions can be meaningfully evaluated (McArdle, 2007; McArdle & Cattell, 1994; Meredith, 1993). This may be particularly important for establishing the clinical utility and interpretability of brief, self-report symptom questionnaires, such as the IES-R and others mentioned here, given the unique usefulness of these measures for tracking changes over time or across treatment through repeated administrations (Weiss, 2004). Conversely, the detection of changes in factor structure over time could offer meaningful insight regarding the development and course of the disorder and signal that further revisions of the PTSD construct and measures used to represent it are necessary.

In summary, prior studies of the factor structure of PTSD have supported the splitting of avoidance and numbing symptoms into separate factors, but consensus on the optimal overall structure has yet to emerge. Two four-factor models have received strong empirical support, a pair of two-factor models have received limited empirical support, and research has suggested that sleep disturbance may constitute a distinct factor for the focal instrument of the current study, the IES-R. Finally, further research on the cross-cultural validity and longitudinal stability of measures of PTSD is needed. To address these issues, the present study examined the suitability of a number of conceptually and empirically informed PTSD factor models using the IES-R. We first evaluated which models seemed to best characterize factor structure in separate cross-sectional analyses for each of two culturally distinct samples that varied with regard to nationality, language version of the IES-R that was administered, and other personal and trauma characteristics. Next, we conducted multiple-Group CFAs to assess the extent to which the preferred models were invariant across the two samples. The best-fitting model then was tested for longitudinal invariance within one of these samples.

Method

Participants

Israeli Emergency Room Sample

This sample (N = 235) included 130 (55%) male and 105 (45%) female medical patients who presented at a hospital emergency room in a large city in Israel. All met screening standards for exposure to a highly stressful PTSD Criterion A event (American Psychiatric Association, 2000), as assessed by a research psychiatrist or psychologist, and persons with certain medical conditions (head injury, history of psychosis, or substance abuse) were excluded. Participants were part of a longitudinal study examining event-related variables, objective psychophysiological reactivity indicators, and measures of mental health outcomes (e.g., Bachar, Hadar, & Shalev, 2005). Most experienced severe motor vehicle accidents (80%), while others experienced terrorist acts (13%), serious work or domestic accidents (3%), or other traumatizing events (4%). Participants ranged in age from 17 to 65 (M = 31.57, SD = 10.72). Participants reported being married (47%), never married (46%), divorced/separated (6%), or widowed (1%). This sample provided IES-R data on up to three occasions, averaging 9.93 days, 39.78 days, and 168.37 days following emergency room admission.

U.S. Undergraduate Student Sample

Participants (N = 306) included 84 (27%) male and 222 (73%) female psychology undergraduate students at a medium-size rural Southern U.S. university who received course credit for their participation. They were part of a larger study examining the role of childhood temperament as a risk factor for subsequent development of PTSD. Trauma exposure was assessed using the Traumatic Events Questionnaire (Lauterbach & Vrana, 1993, 1996; Vrana & Lauterbach, 1994). The majority of the sample (88%) reported experiencing at least one traumatic event. The most frequently reported events were learning of the sudden or unexpected death of a loved one (54%), serious accident (46%), natural disaster (36%), and otherwise being in a dangerous or life-threatening situation (15%).[2] Participants ranged in age from 17 to 56 years (M = 21.76, SD = 5.96). The majority (77%) was single; 18% were married or living with a partner, and 6% were divorced or separated. Most participants (65%) identified as European American, although a substantial minority (32%) identified as African American, and a small number (3%) identified other ethnic backgrounds. This undergraduate student sample provided data on one occasion; the average time since exposure to the focal ("worst") traumatic event was 6.40 years, with a portion of the group (6%) reporting ongoing trauma at data collection.

Measure

IES-R

Posttraumatic stress symptoms were assessed in both samples using the IES-R (Weiss & Marmar, 1997), a 22-item self-report questionnaire measuring symptoms frequently endorsed following exposure to a distressing event. The IES-R is a revised version of the original IES (Horowitz et al., 1979). The most significant content revision was the addition of items assessing hyperarousal to reflect *DSM* cluster D symptoms. In addition, one item was added to capture dissociative aspects of flashbacks. Finally, a sleep item inquiring about problems with sleep onset or maintenance was split into two separate items, to address the common recommendation to avoid "double barreled" questions (e.g., Clark & Watson, 1995); "trouble falling asleep" was determined to index hyperarousal, and "trouble staying asleep" was assigned to the intrusion category. As a result of these revisions, the IES-R includes items representing all three *DSM–IV* (American Psychiatric Association, 2000) PTSD symptom clusters. As previously noted, unlike many commonly used self-report or interview-based measures, the IES-R items do not map rigidly to the 17 *DSM–IV* PTSD symptoms, yet IES-R item content is faithful to the primary criteria.

The IES-R has demonstrated strong psychometric properties in prior research. The internal consistency reliability of the full scale has been reported as .96 (Creamer et al., 2003). Excellent internal consistency reliability estimates for the subscales also have been reported, ranging from the low .80s to low .90s (Weiss, 2004; Weiss & Marmar, 1997). Convergent validity of the IES-R has been demonstrated through associations with a number of other PTSD measures (Corapcioglu, Yargic, Geyran, & Kocabasglu, 2006; Erickson & Steiner, 2000; Paunovic &

[2] Total percentage exceeds 100 because many respondents reported more than one traumatic event.

Ost, 2005; Shapinsky, Rapport, Henderson, & Axelrod, 2005).

Participants in the emergency room and undergraduate samples completed versions of the IES-R that differed in two regards. First, the emergency room sample completed a Hebrew-language version (subjected to the expected standards of translation and back-translation by bilingual professionals; Shalev & Freedman, 2005), whereas the undergraduate sample completed the measure in its original English format. Second, the instructions and response anchors differed across samples. For the emergency room sample, the instructions and anchors comprised those that were used for the original IES and in initial validation research of the IES-R; participants were instructed to rate "how frequently" each item occurred during the previous seven days on a 4-point response scale (0 = *not at all*, 1 = *rarely*, 3 = *sometimes*, and 5 = *often*). The version completed by the undergraduate sample, on the other hand, included the revised anchors and instructions later recommended for use by Weiss and Marmar (1997); participants were instructed to rate how much each item "bothered or distressed" them during the previous seven days on a 5-point scale (0 = *not at all,* 1 = *a little bit,* 2 = *moderately,* 3 = *quite a bit,* and 4 = *extremely*). Participants who had endorsed traumatic events on the Traumatic Events Questionnaire were instructed to reference the "worst event experienced"; participants who did not endorse a specific event on this instrument were requested to identify (in writing) the most distressful event they had experienced in their lives and reference that stressor in completing the IES-R.

Overview of Data Analyses

Descriptive statistics and estimates of scale reliability for each sample and total scale intercorrelations for the repeated assessments of the emergency room sample first were calculated. Next, a series of factor models (see Table 1) was evaluated separately for each sample; for the emergency room sample, data from the Time 1 IES-R administration were used. These models included a one-factor global PTSD model (Model 1; as a reference point); a two-

Table 1
Competing Factor Models

	Model								
Item	1	2A	2B	3	4A	4B	4C	5A	5B
1 . . . reminder brought back feelings	P	I-A	P	I	I	I	I	I	I
2 . . . trouble staying asleep	P	I-A	P	I	I	I	S	S	S
3 . . . making me think about it	P	I-A	P	I	I	I	I	I	I
6 . . . thought . . . when I didn't mean to	P	I-A	P	I	I	I	I	I	I
9 . . . popped into my mind	P	I-A	P	I	I	I	I	I	I
14 . . . acting or feeling as though	P	I-A	P	I	I	I	I	I	I
16 . . . waves of strong feelings	P	I-A	P	I	I	I	I	I	I
19 . . . reminders . . . physical reactions	P	I-A	P	I	I	I	I	I	I
20 . . . had dreams about it	P	I-A	P	I	I	I	S	S	S
5 . . . avoided letting myself get upset	P	I-A	P	A-N	A	A	A-N	A	A
8 . . . stayed away from reminders	P	I-A	P	A-N	A	A	A-N	A	A
11 . . . tried not to think about it	P	I-A	P	A-N	A	A	A-N	A	A
17 . . . remove it from memory	P	I-A	P	A-N	A	A	A-N	A	A
22 . . . tried not to talk about it	P	I-A	P	A-N	A	A	A-N	A	A
7 . . . hadn't happened or wasn't real	P	N-H	D	A-N	N	D	A-N	N	D
12 . . . didn't deal with them	P	N-H	D	A-N	N	D	A-N	N	D
13 . . . feelings . . . were kind of numb	P	N-H	D	A-N	N	D	A-N	N	D
4 . . . irritable or angry	P	N-H	D	H	H	D	H	H	D
10 . . . jumpy and easily startled	P	N-H	P	H	H	H	H	H	H
15 . . . trouble falling asleep	P	N-H	D	H	H	D	S	S	S
18 . . . trouble concentrating	P	N-H	D	H	H	D	H	H	D
21 . . . watchful or on-guard	P	N-H	P	H	H	H	H	H	H

Note. P = posttraumatic stress disorder; I-A = Intrusion-Avoidance; N-H = Numbing-Hyperarousal; D = Dysphoria; I = Intrusion; A-N = Avoidance-Numbing; H = Hyperarousal; A = Avoidance; N = Numbing; S = Sleep.

factor numbing model (Model 2A; Taylor et al., 1998), a two-factor dysphoria model (Model 2B; Buckley et al., 1998; Taylor et al., 1998), a three-factor *DSM–IV* model (Model 3; American Psychiatric Association, 2000; Cordova, Studts, Hann, Jacobsen, & Andrykowski, 2000), a four-factor numbing model (Model 4A; King et al., 1998), and a four-factor dysphoria model (Model 4B; Simms et al., 2002). In addition, a distinct sleep factor was incorporated into Models 3, 4A, and 4B to create a four-factor *DSM–IV* model with sleep (Model 4C), a five-factor numbing model with sleep (Model 5A), and a five-factor dysphoria model with sleep (Model 5B), respectively. The best-fitting models for each sample were identified, and the number of factors and pattern of loadings were examined to judge the degree of invariance across samples.

Using the preferred models from these cross-sectional analyses, invariance next was assessed in terms of the goodness of fit of a multiple-group factor model, where the same number of factors and pattern of factor loadings were specified for both groups. Because of the aforementioned differences in raw score item responses, only configural invariance could be considered. Finally, the best-fitting model was used to gauge invariance over time for the longitudinal emergency room data. Here, we followed recommendations set forth by McArdle (2007), wherein metric invariance over time is demonstrated by a well-fitting multilevel model having loadings for a pooled within-occasions structure equal to loadings for a structure underlying item scores averaged over occasions.

For all analyses, a robust maximum likelihood estimator was employed, and mean and covariance structures of item-level data were analyzed. Items were specified to load on a single factor, and covariances among item residuals were fixed at zero, the latter constraint imposed in line with the standard assumption of independent and identically distributed error (Browne & Nesselroade, 2005; McArdle, 2007; Meredith & Horn, 2001). Indices used to aid in model selection included the comparative fix index (CFI; Bentler, 1990), the Tucker-Lewis or nonnormed fit index (TLI; Bentler & Bonett, 1980), the root mean square error of approximation (RMSEA; Steiger, 1990), the standardized root-mean-square residual (SRMR; Bentler, 1990), and the Bayesian information criterion (BIC; Schwarz, 1978). The CFI and TLI compare the fit of the proposed model to one in which all item covariances are fixed at zero, and higher values are considered good (>.90, acceptable, and >.95, desirable; Hu & Bentler, 1998). The RMSEA is an index of misfit per degree of freedom; lower values are preferred (<.08, acceptable, <.06 desirable; Hu & Bentler, 1998). The SRMR is the average standardized deviation in the model-based reproduced covariances in contrast to those observed in the data; lower values are optimal (<.08; Hu & Bentler, 1998). For the BIC, no firm standards for values are available, but this fit index is used to compare among rival or competing models, even those that are not nested, with values closer to zero preferred. Jeffreys (1961) detailed original guidelines for interpretation of BIC differences for competing models, which were later updated by Raftery (1995): A difference of 6–10 indicates strong support and >10 indicates very strong or decisive support for the model with the lower (absolute) BIC value. All model testing was accomplished using the Mplus (version 3) software package (Muthén & Muthén, 1998–2005). Additional details on data analyses are integrated into the Results section to follow.

Results

Descriptive and Summary Statistics

Table 2 presents the mean, *SD*, and range for each IES-R administration to the Israeli emergency room sample (Time 1, Time 2, and Time 3), as well as the same statistics for the single IES-R administration to the U.S. undergraduate student sample. As one might expect, as time progressed, the average IES-R score for the emergency room sample decreased. At all three occasions, the scores demonstrated a good deal of dispersion, as represented by *SD* and range, with relatively little change in variability over time. The values of the *SD*s relative to the means suggest some degree of positive skew. For this sample, the internal consistency reliability remained strong over time. IES-R scores for the single administration to undergraduates also indicate good dispersion, a positively skewed distribution, and solid internal consistency reliability. Direct comparison of means between samples is not possible, owing to their different anchoring and scoring methods

Table 2
Descriptive and Summary Statistics for Impact of Event Scale-Revised

Sample	*M*	*SD*	Range	α[a]	Intercorrelations	
					ERT1	ERT2
Emergency room sample Time 1	45.03	25.52	0–106	.93		
Emergency room sample Time 2	32.66	24.31	0–99	.95	.82	
Emergency room sample Time 3	28.45	23.73	0–96	.94	.66	.80
Undergraduate sample	20.16	19.82	0–82	.96		

Note. Means, *SD*s, and intercorrelations among emergency room sample scores are maximum likelihood estimates, given the available data ($N = 235$). ERT1 = emergency room sample Time 1; ERT2 = emergency room sample Time 2.
[a] Internal consistency reliability calculated as Cronbach's alpha.

for raw scores. The rightmost columns of Table 2 display the three correlation coefficients among scores for sequential administrations of the IES-R to the emergency room sample. The pattern is consistent with the Markov simplex, wherein associations between adjacent assessments are stronger than that between more distal assessments.

Factor Structure Over Groups

Table 3 contains results for the nine factor models specified and tested separately for each of the two samples. As shown there, for both samples, the three models having a separate sleep factor—the four-factor *DSM–IV* model with sleep (4C) and the two five-factor models (5A and 5B), numbing and dysphoria, respectively, each with a sleep factor—provided markedly better fit to the data than their counterparts that did not incorporate a separate sleep factor (Models 3, 4A, and 4B). Furthermore, Table 3's collection of fit indices suggests that Model 5A provided a better fit to the data than Model 5B across both samples. Nonetheless, for the emergency room sample, the more parsimo-

Table 3
Cross-Sectional Confirmatory Factor Analysis Results for Separate Samples

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR	BIC
Emergency room sample (Time 1)							
1 (one-factor global PTSD)	556.47	209	.81	.79	.097	.070	898.10
2A (two-factor numbing)	553.92	208	.81	.79	.097	.071	900.73
2B (two-factor dysphoria)	554.16	208	.81	.79	.097	.070	900.97
3 (three-factor *DSM-IV*)	434.97	206	.87	.86	.079	.059	792.13
4A (four-factor numbing)	434.69	203	.87	.85	.080	.059	807.38
4B (four-factor dysphoria)	453.62	203	.86	.84	.084	.062	826.31
4C (four-factor *DSM-IV* with sleep)[a]	340.55	203	.92	.91	.062	.062	713.24
5A (five-factor numbing with sleep)[a]	337.30	199	.92	.91	.063	.062	730.69
5B (five-factor dysphoria with sleep)	350.14	199	.92	.90	.066	.064	743.53
Undergraduate sample							
1 (one-factor global PTSD)	532.60	209	.88	.87	.073	.055	907.49
2A (two-factor numbing)	510.09	208	.89	.88	.070	.055	890.66
2B (two-factor dysphoria)	522.90	208	.89	.87	.072	.055	903.47
3 (three-factor *DSM-IV*)	382.19	206	.94	.93	.054	.047	774.12
4A (four-factor numbing)	368.33	203	.94	.93	.053	.046	777.30
4B (four-factor dysphoria)	398.60	203	.93	.92	.057	.048	807.57
4C (four-factor *DSM-IV* with sleep)[b]	327.59	203	.96	.95	.046	.044	736.56
5A (five-factor numbing with sleep)[b]	313.29	199	.96	.95	.044	.043	744.98
5B (five-factor dysphoria with sleep)	333.65	199	.95	.94	.048	.045	765.34

Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; BIC = Bayesian information criterion; PTSD = posttraumatic stress disorder; *DSM-IV* = *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*.
[a] Model 4C versus Model 5A: Corrected $\Delta\chi^2(4) = 3.08$, $p = .54$. [b] Model 4C versus Model 5A: Corrected $\Delta\chi^2(4) = 13.24$, $p = .01$.

nious Model 4C can be judged superior to Model 5A, as reflected in a smaller RMSEA and in the chi-square difference between these models, corrected $\Delta\chi^2$ (4) = 3.08, p = .54. It is important to note for this sample, Model 4C's BIC value is 17.45 lower than Model 5A's BIC value, providing what Raftery (1995) characterizes as very strong evidence that Model 4C (over Model 5A) produced the data, with odds of at least 150:1 in its favor. For the undergraduate sample, the RMSEA and SRMR values for Model 5A are smaller than those for Model 4C, and the corrected $\Delta\chi^2$ (4) = 13.24, p = .01. But, the difference in BIC values is 8.42, translated to strong evidence in support of Model 4C (over Model 5A), with odds of greater than 60:1 in its favor (Raftery, 1995).

For completeness of comparison, Models 4C, 5A, and 5B were retained for the multiple-group analyses regarding configural invariance of the factor structure of PTSD as measured by the IES-R when both the Israeli emergency room (Time 1 data) and U.S. undergraduate samples were treated simultaneously. For each of these three PTSD structures, intercepts and loadings were unconstrained or free over groups. As shown in Table 4, the indices of model-data fit are generally good for all three models, with Models 4C and 5A again favored over Model 5B. However, scrutiny of differences in BIC values over the three models argues for the preeminence of Model 4C. Despite a slightly larger SRMR for Model 4C (compared to Model 5A), the BIC difference of 29.79 decisively (to use the language of Jeffreys, 1961) or very strongly (returning to the grades of evidence offered by Raftery, 1995) endorses the choice of Model 4C for the two groups.

Factor Structure Over Time

The examination of factorial invariance over three assessments focused on the accepted model from prior analyses, Model 4C (four-factor *DSM–IV* with sleep). Table 5 presents the results of these multilevel analyses using longitudinal data from the Israeli emergency room sample. Borrowing from the logic of the analysis of variance, McArdle (2007) reasoned that configural invariance would be supported by a well-fitting model that specified structures both between and within occasions having the same number of factors and pattern of loadings, covariances, and residuals. Furthermore, metric invariance would be upheld if equality constraints on the loadings for both the between- and within-occasions components yielded acceptable fit to the data.

The first row of Table 5 ("Fully [all factor loadings] free") provides support for configural invariance; all fit indices are relatively strong in the context of contemporaneous standards (Hu & Bentler, 1998). The second row ("Fully [all factor loadings] fixed") provides results when factor loadings for the between- and within-occasions components were constrained to be equal. Compared to the fully free findings, constraints on the loadings appear to considerably damage model-data fit, corrected $\Delta\chi^2$ (16) = 71.87, $p < .001$.

To determine the possible source of the instability of loadings over assessment occasions, we systematically freed loadings for one factor at a time. If freeing a factor's equality constraints on the between- and within-occasions components yielded a dramatically better fit to the data, then loadings on that factor could not be considered equivalent, and hence would be the source of instability. Results are displayed in the remaining rows of Table 5. The most striking finding relates to freeing the loadings for the intrusion factor, which leads to a marked improvement in fit compared to the model with fully fixed loadings. Here, corrected $\Delta\chi^2$ (6) = 60.48, $p < .001$, indicating that the loadings on the intrusion factor vary over time.

Table 4
Multiple-Group Confirmatory Factor Analysis Results

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR	BIC
Model 4C	655.80	406	.95	.94	.052	.052	1551.79
Model 5A	646.30	398	.95	.94	.052	.051	1581.52
Model 5B	680.85	398	.94	.93	.055	.053	1616.07

Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; BIC = Bayesian information criterion.

Table 5
Longitudinal Confirmatory Factor Analysis Results for Model 4C

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR between	SRMR within	BIC
Fully (all factor loadings) free	537.95	328	.95	.94	.035	.064	.061	35013.53
Fully (all factor loadings) fixed	634.33	344	.93	.92	.040	.080	.059	35035.95
Intrusion factor loadings free	550.58	338	.95	.94	.035	.069	.064	34976.09
Avoidance-Numbing factor loadings free	643.31	338	.93	.92	.042	.085	.056	35068.18
Hyperarousal factor loadings free	623.33	341	.93	.93	.040	.076	.056	35041.31
Sleep factor loadings free	634.76	343	.93	.92	.040	.080	.059	35042.04

Note. CFA = confirmatory factor analysis; χ^2 = chi-square; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; BIC = Bayesian information criterion.

Additionally, the BIC, reflecting not only goodness of fit but also parsimony, favors the more constrained intrusion free model over its fully free or most saturated counterpart. The difference in BIC values reported for the fully free model and the intrusion free model is a striking 37.44 points.

It is important to note that the findings presented in Table 5 are based on 20 of the total of 22 items in the IES-R. The reason for removing two items was that the reproduced information matrix for the between-occasions component was not positive definite. Upon closer inspection, the source of the problem was extraordinarily high collinearity over occasions between Item 2 (". . . trouble staying asleep") and Item 15 (". . . trouble falling asleep"), as well as between Item 11 (". . . tried not to think about it") and Item 22 (". . . tried not to talk about it"). When Items 15 and 22 were removed, convergence to a solution was achieved, presumably with little loss to the integrity of the instrument's content. To verify this latter point, the previously presented multiple-group analyses were reconducted with the 20 items, and the pattern of findings mirrored that reported in Table 4.

Discussion

This study was aimed at determining the most parsimonious and veridical representation of the structure of PTSD as measured by the IES-R, and demonstrating evidence for invariance of the best-fitting structure over both groups and occasions. Three models, each of which specified a separate sleep factor, yielded fit indices that were acceptable to very good. One of these, Model 4C (four-factor *DSM–IV* model with sleep), was judged to provide the best fit to the data. That is, when single-occasion data from two distinct samples were analyzed separately, this model garnered initial support for configural invariance (a common set of factors with a common pattern of item loadings). With the same single-occasion data subjected to multiple-group analyses, configural invariance for Model 4C was affirmed. Configural invariance for this model was also upheld when longitudinal data from one sample were analyzed in a multilevel model to ascertain stability across time. Evidence for metric invariance was qualified: While loadings on avoidance-numbing, hyperarousal, and sleep appeared stable over occasions, loadings on the intrusion factor did not.

Consistent with prior inquiry into the structure of PTSD using other measures (e.g., Palmieri et al., 2007), the one- and two-factor solutions (Models 1, 2A, and 2B) for both samples proved weak in contrast to other models (see Table 3). Yet, though not acceptable, the fits of the three-factor *DSM–IV* solution (Model 3) versus the four-factor numbing and dysphoria solutions (Models 4A and 4B, respectively) observed here are far more comparable to one another than comparisons between three- and four-factor solutions in previous work (e.g., Palmieri et al., 2007; Simms et al., 2002). In fact, with regard to Models 3 and 4A for both samples, the normed fit indices (CFI and TLI) as well as the RMSEA and SRMR are close in value; only the BIC clearly penalizes the four-factor numbing model for its three additional estimated factor covariances. Moreover, all fit indices for both samples would favor the more parsimonious *DSM–IV* solution (Model 3) to the four-factor dysphoria solution (Model 4B). Overall, the three- and four-factor models for the undergraduate sample yielded reasonably good fit to the data, with CFI and TLI in the low- to mid-.90s and the RMSEA and SRMR within acceptable limits. That being said, neither the three- nor the four-factor numbing and dysphoria models provided as strong a fit to the data as Models 4C, 5A, and 5B. As shown in Table 3, for both samples, there is marked improvement in model-data fit when the IES-R sleep items are specified to be indicators of a separate factor. Furthermore, the collection of fit indices for Model 4C (vs. Models 5A and 5B) do not make a strong case for disaggregating the avoidance-numbing item set into separate factors for this particular measure.

The sleep factor that emerged in these analyses is reminiscent of findings by Larsson (2000) for IES data from persons who witnessed a mass murder and Amdur and Liberzon (2001) for IES data from a combat veteran sample. Two of the three items that formed the sleep factor were drawn from the IES-R intrusion subscale (“I had dreams about it” and “I had trouble staying asleep”), and the third was drawn from the IES-R hyperarousal subscale (“I had trouble falling asleep”). These three items mirror closely two *DSM–IV* criteria, recurrent distressing dreams and difficulty falling and staying asleep. Future researchers examining the structure of PTSD using the IES-R or other instruments should certainly consider a test of such a factor. Moreover, the findings highlight the importance of targeting sleep disturbance as perhaps deserving special consideration in understanding the pathogenesis and nature of PTSD and in structuring and evaluating treatment interventions (see important contributions to the sleep-PTSD literature by Davis & Wright, 2007; Harvey, Jones, & Schmidt, 2003; Krakow et al., 2001; Mellman, 1997; Ross et al., 1994; Woodward, Murburg, & Bliwise, 2000). One might also entertain the possibility that this cluster of sleep problems is a collection of nonspecific symptoms that PTSD shares with other mood and anxiety disorders, paralleling the Simms et al. dysphoria conceptualization.

Why might the avoidance and numbing aspects of PTSD as measured by the IES-R not clearly replicate the fairly well-accepted notion that they are distinct though related PTSD factors (e.g., King et al., 2006)? Careful scrutiny of the *DSM–IV–TR* (American Psychiatric Association, 2000) criteria, as well as the content of other instruments that closely mirror these criteria, vis-à-vis the IES-R item content, provides a possible explanation. *DSM–IV* Criteria C1 and C2 make direct reference to active avoidance of thoughts, feelings, conversations, activities, places, and people that can trigger recollections of the traumatic event; these criteria are definitely tied to the event. On the other hand, with the exception of Criterion C3 (psychogenic amnesia for the event), the set of emotional numbing Criterion C symptoms (C4 - C7), describe diminished affect and reduced interactions with others, social withdrawal, and retreat from the normal expectations of a fulfilling life, none of which reference feelings specifically tied to the traumatic event but rather connote an overall state of malaise and distress. Therefore, in analysis of PTSD measures with items paralleling the *DSM–IV* criteria, it is rather straightforward to expect and detect two factors, one reflecting direct attempts to fend off potentially intrusive stimuli linked to a prior trauma (active avoidance), and the second a general blunting or deadening of affect and unwillingness to engage one’s surroundings (emotional numbing). And, the empirical literature to date endorses this rationale.

Turning to the IES-R, it is noteworthy that all eight avoidance-numbing items specifically reference the event in their item statements, even the item that most directly addresses emotional numbing (“My feelings about it [the event] were kind of numb”). Unlike *DSM–IV* and PTSD measures with items mapped strictly to the 17 *DSM–IV* symptoms, the IES-R does not overtly assess feelings of detachment from others, isolation, or other facets of intimacy or relationship quality. Moreover, its eight avoidance-numbing items largely appear to be saturated with content suggesting active avoidance of reminders of the event. If, indeed, emotional numbing as operationalized by Criteria C4–C7 (e.g., withdrawal from others, inability to form meaningful relationships) is a valued and valid aspect of PTSD, then perhaps the IES-R could benefit from an additional core set of items with content that addresses this aspect with greater fidelity.

Despite the differences in samples (nationality, language, trauma characteristics, and time since exposure), the multiple-group results (see

Table 4) argue fairly strongly for configural invariance for the underlying structure of the IES-R. Thus, the contribution of the Model 4C factors to the observed responses on the IES-R could be judged stable across the two samples, makes a case for the cross-sectional equivalence of measurement across instrument versions, and speaks to the cross-cultural validity of the PTSD construct as reflected in IES-R item content. The findings also may support the contention of Elhai et al. (2006) that there is negligible incremental information between assessments that index frequency versus intensity/severity of PTSD symptoms.

The portion of this study that examined the invariance of factor structure over time (see Table 5) revealed stability in the number of factors and pattern of loadings over three occasions assessed on average at approximately 10 days, 40 days, and 168 days following trauma exposure. On the other hand, respondents did not necessarily perceive the intrusion items as having the same meaning from occasion to occasion subsequent to exposure to the traumatic event. Of interest, the particular items for which there was a conspicuous discrepancy in loadings from the between-occasions to the within-occasions solutions, key to the potential locus of the instability, were those that describe the more dramatic or intense intrusive responses ("I found myself acting or feeling as though I was back at that time," "I had waves of strong feelings about it," and "Reminders of it caused me to have physical reactions . . .").[3] Perhaps, over time, as many trauma victims engage in processes of meaning-making, cognitive restructuring, and/or habituation, either naturally or in the therapeutic context, at least some are able to resolve the more profound coping difficulties reflected in these items. Hence, the representation of the event, and the perceptions of the actual item statements about the event, differentially shift, thereby leading to variation in the item loadings on the intrusion factor from occasion to occasion. Future research should attempt to replicate this finding of change over time in the associations between the intrusion factor and its manifest item indicators and clarify explanations for the observed shift in construct meaning. Users of the IES-R also should be aware that any longitudinal analysis of IES-R data at the factor level mandates caution in assuming the intrusion component of PTSD is equivalent from occasion to occasion.

There are limitations to this study. As previously described, the measures administered to the separate samples differed in language, instructional set, and item response formats. Although the Israeli emergency room sample was comprised of persons who had a verified *DSM–IV* Criterion A life threatening event with concurrent medical emergency, the U.S. undergraduate sample self-reported exposure and had a mixed trauma history surveyed across the life span. For this sample, there was less scrutiny to the intensity and circumstances of stressor events, and thus there remains ambiguity concerning the extent of exposure typically required for a PTSD diagnosis. Despite these differences in data and sample characteristics, however, the multiple-group results argue fairly strongly for configural invariance of the structure of PTSD. Additionally, we duly recognize that the removal of two items to accommodate convergence in the longitudinal factor analysis might temper conclusions concerning the stability of factor structure over time.

In summary and in closing, these factor analyses, with specific attention to invariance over samples and time, offer three substantive conclusions regarding the structure of PTSD, as measured by the IES-R, an instrument used quite broadly around the world. First, they argue for consideration of a separate sleep factor and perhaps heightened attention to sleep dysfunction within the assessment of PTSD. Second, the strong showing of Model 4C suggests that the overwhelming evidence to date for splitting avoidance symptoms from numbing symptoms does not necessarily hold for this particular instrument. The caveat here is that numbing in the IES-R does not appear to carry exactly the same meaning as that of the *DSM–IV* and other instruments with item content strictly faithful to the *DSM–IV* criteria. Third, the study indicated that intrusion may not be a stable PTSD component following exposure to a traumatic event. Recollection of the traumatic event in terms of intrusive thoughts, feelings, dreams, and phys-

[3] Full results for all models discussed in this article, including additional details on factor loadings, intercepts, variances, and covariances, and residual variances may be obtained from the first author.

ical reactions possibly differ as a function of time since the event, prompting an additional avenue of research inquiry and vigilance in using the IES-R to assess change in this component of PTSD.

References

Amdur, R. L., & Liberzon, I. (2001). The structure of posttraumatic stress disorder symptoms in combat veterans: A confirmatory factor analysis of the Impact of Event Scale. *Journal of Anxiety Disorders, 15,* 345–357.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Asmundson, G. J. G., Wright, K. D., McCreary, D. R., & Pedlar, D. (2003). Post-traumatic stress disorder symptoms in United Nations peacekeepers: An examination of factor structure in peacekeepers with and without chronic pain. *Cognitive Behaviour Therapy, 32,* 26–37.

Bachar, E., Hadar, H., & Shalev, A. Y. (2005). Narcissistic vulnerability and the development of PTSD: A prospective study. *Journal of Nervous and Mental Disease, 193,* 762–765.

Baschnagel, J. S., O'Connor, R. M., Colder, C. R., & Hawk, L. W., Jr. (2005). Factor structure of posttraumatic stress among Western New York undergraduates following September 11th terrorist attack on the World Trade Center. *Journal of Traumatic Stress, 18,* 677–684.

Beck, J. G., Grant, D. M., Read, J. P., Clapp, J. D., Coffey, S. F., Miller, L. M., et al. (2008). The Impact of Event Scale–Revised: Psychometric properties in a sample of motor vehicle accident survivors. *Journal of Anxiety Disorders, 22,* 187–198.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., et al. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress, 8,* 75–90.

Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of ARMA time series models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A Festschrift for Roderick P. McDonald* (pp. 415–452). Mahwah, NJ: Erlbaum.

Buckley, T. C., Blanchard, E. B., & Hickling, E. G. (1998). A confirmatory factor analysis of posttraumatic stress symptoms. *Behavior Research and Therapy, 36,* 1091–1099.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7,* 309–319.

Corapcioglu, A., Yargic, I., Geyran, P., & Kocabasglu, N. (2006). Validity and reliability of a Turkish version of the Impact of Event Scale-Revised (IES-R). *Yeni Symposium, 44,* 14–22.

Cordova, M. J., Studts, J. L., Hann, D. M., Jacobsen, P. B., & Andrykowski, M. A. (2000). Symptom structure of PTSD following breast cancer. *Journal of Traumatic Stress, 13,* 301–319.

Creamer, M., Bell, R., & Failla, S. (2003). Psychometric properties of the Impact of Event Scale-Revised. *Behavior Research and Therapy, 41,* 1489–1496.

Davis, J. L., & Wright, D. C. (2007). Randomized clinical trial for treatment of chronic nightmares in trauma-exposed adults. *Journal of Traumatic Stress, 20,* 123–133.

DuHamel, K. N., Ostroff, J., Ashman, T., Winkel, G., Mundy, E. A., Keane, T. M., et al. (2004). Construct validity of the Posttraumatic Stress Disorder Checklist in cancer survivors: Analyses based on two samples. *Psychological Assessment, 16,* 255–266.

Elhai, J. D., Lindsay, B. M., Gray, M. J., Grubaugh, A. L., North, T. C., & Frueh, B. C. (2006). Examining the uniqueness of frequency and intensity symptom ratings in posttraumatic stress disorder assessment. *Journal of Nervous and Mental Disease, 194,* 940–944.

Erickson, S. J., & Steiner, H. (2000). Somatic symptoms in long-term pediatric cancer survivors. *Psychosomatics, 41,* 339–346.

Foa, E. B., Cashman, L. A., Jaycox, L., & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological Assessment, 9,* 445–451.

Foa, E. B., Riggs, D. S., & Gershuny, B. S. (1995). Arousal, numbing, and intrusion: Symptom structure of PTSD following assault. *American Journal of Psychiatry, 152,* 116–120.

Harvey, A. G., Jones, C., & Schmidt, D. A. (2003). Sleep and posttraumatic stress disorder: A review. *Clinical Psychology Review, 23,* 377–407.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. *Psychosomatic Medicine, 41,* 209–218.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453.

Jeffreys, H. (1961). *Theories of probability* (3rd ed.). Oxford: Oxford University Press.

Joseph, S. (2000). Psychometric evaluation of Horowitz's Impact of Event Scale: A review. *Journal of Traumatic Stress, 13,* 2000.

King, D. W., Leskin, G. A., King, L. A., & Weathers, F. A. (1998). Confirmatory factor analysis of the Clinician-Administered PTSD Scale: Evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment, 10,* 90–96.

King, L. A., & King, D. W. (1994). Latent structure of the Mississippi Scale for combat-related posttraumatic stress disorder: Exploratory and higher-order confirmatory factor analysis. *Assessment, 1,* 275–291.

King, L. A., King, D. W., Orazem, R. J., & Palmieri, P. A. (2006). Research on the latent structure of PTSD. *PTSD Research Quarterly, 17,* 1–7.

Krakow, B., Hollifield, M., Johnston, L., Koss, M., Schrader, R., Warner, T. D., et al. (2001). Imagery rehearsal therapy for chronic nightmares in sexual assault survivors with posttraumatic stress disorder: A randomized controlled trial. *Journal of the American Medical Association, 286,* 537–545.

Krause, E. D., Kaltman, S., Goodman, L. A., & Dutton, M. A. (2007). Longitudinal factor structure of posttraumatic stress symptoms related to intimate partner violence. *Psychological Assessment, 19,* 165–175.

Larsson, G. (2000). Dimensional analysis of the Impact of Event Scale using structural equation modeling. *Journal of Traumatic Stress, 13,* 2000.

Lauterbach, D., & Vrana, S. (1993, October). *Validation of a measure of PTSD among college students.* Paper presented at the International Society for Traumatic Stress Studies, San Antonio, TX.

Lauterbach, D., & Vrana, S. R. (1996). Three studies on the reliability and validity of a self-report measure of posttraumatic stress disorder. *Assessment, 3,* 17–25.

Marshall, G. N. (2004). Posttraumatic Stress Disorder Symptom Checklist: Factor structure and English-Spanish measurement invariance. *Journal of Traumatic Stress, 17,* 223–230.

McArdle, J. J. (2007). Five steps in the factor analysis of longitudinal data. In R. Cudeck & R. MacCallum (Eds.), *Factor analysis at* 100 (pp. 99–130). Mahwah, NJ: Erlbaum.

McArdle, J. J., & Cattell, R. B. (1994). Structural equation models of factorial invariance in parallel proportional profiles and oblique confactor problems. *Multivariate Behavioral Research, 29,* 63–113.

McWilliams, L. A., Cox, B. J., & Asmundson, G. J. G. (2005). Symptom structure of posttraumatic stress disorder in a nationally representative sample. *Journal of Anxiety Disorders, 19,* 626–641.

Mellman, T. (1997). Psychobiology of sleep disturbances in posttraumatic stress disorder. *Annals of the New York Academy of Sciences, 821,* 142–149.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525–543.

Meredith, W., & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 203–240). Washington, DC: American Psychological Association.

Muthén, B., & Muthén, L. (1998–2005). *Mplus user's guide.* Los Angeles: Author.

Norris, F. H., & Hamblen, J. L. (2004). Standardized self-report measures of civilian trauma and PTSD. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (2nd ed., pp. 63–102). New York: Guilford Press.

Norris, F. H., & Perilla, J. L. (1996). The Revised Civilian Mississippi Scale for PTSD: Reliability, validity, and cross-language stability. *Journal of Traumatic Stress, 9,* 285–298.

Norris, F. H., Perilla, J. L., & Murphy, A. D. (2001). Postdisaster stress in the United States and Mexico: A cross-cultural test of the multicriterion conceptual model of posttraumatic stress disorder. *Journal of Abnormal Psychology, 110,* 553–563.

Palmieri, P. A., & Fitzgerald, L. F. (2005). Confirmatory factor analysis of posttraumatic stress symptoms in sexually harassed women. *Journal of Traumatic Stress, 18,* 657–666.

Palmieri, P. A., Weathers, F. W., Difede, J., & King, D. W. (2007). Confirmatory factor analysis of the PTSD Checklist and the Clinician-Administered PTSD Scale in disaster workers exposed to the World Trade Center Ground Zero. *Journal of Abnormal Psychology, 116,* 329–341.

Paunovic, N., & Ost, L. (2005). Psychometric properties of a Swedish translation of the Clinician-Administered PTSD Scale-Diagnostic Version. *Journal of Traumatic Stress, 18,* 161–164.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25,* 111–163.

Rasmussen, A., Smith, H., & Keller, A. S. (2007). Factor structure of PTSD symptoms among West and Central African refugees. *Journal of Traumatic Stress, 20,* 271–280.

Ross, R. J., Ball, W. A., Dinges, D. F., Kribbs, N. B., Morrison, A. R., Silver, S. M. et al. (1994). Rapid eye movement sleep disturbance in posttraumatic stress disorder. *Biological Psychiatry, 35,* 195–202.

Sack, W. H., Seeley, J. R., & Clarke, G. N. (1997). Does PTSD transcend cultural barriers? A study from the Khmer Adolescent Refugee Project. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 49–54.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Shalev, A. Y., & Freedman, S. (2005). PTSD following terrorist attacks: A prospective evaluation. *American Journal of Psychiatry, 162,* 1188–1191.

Shapinsky, A. C., Rapport, L. J., Henderson, M. J., & Axedrod, B. N. (2005). Civilian PTSD scales. *Assessment, 12,* 220–230.

Shelby, R. A., Golden-Kreutz, D. M., & Andersen, B. L. (2005). Mismatch of posttraumatic stress disorder (PTSD) symptoms and *DSM–IV* symptom clusters in a cancer sample: Exploratory factor analysis of the PTSD Checklist-Civilian version. *Journal of Traumatic Stress, 18,* 347–357.

Shevlin, M., Hunt, N., & Robbins, I. (2000). A confirmatory factor analysis of the Impact of Event Scale using a sample of World War II and Korean War veterans. *Psychological Assessment, 12,* 414–417.

Simms, L. J., Watson, D., & Doebbeling, B. N. (2002). Confirmatory factor analysis of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology, 111,* 637–647.

Steiger, J. H. (1990). Structural model evaluation and modification: An internal estimation approach. *Multivariate Behavioral Research, 25,* 173–180.

Taylor, S., Kuch, K., Koch, W. J., Crockett, D. J., & Passey, G. (1998). The structure of posttraumatic stress symptoms. *Journal of Abnormal Psychology, 107,* 154–160.

van der Ploeg, E., Mooren, T. T. M., Kleber, R. J., van der Velden, P. G., & Brom, D. (2004). Construct validation of the Dutch version of the Impact of Even Scale. *Psychological Assessment, 16,* 16–26.

Vrana, S. R., & Lauterbach, D. (1994). Prevalence of traumatic events and post-traumatic psychological symptoms in a nonclinical sample of college students. *Journal of Traumatic Stress, 7,* 289–302.

Weathers, F. W., Litz, B. T., Herman, D. W., Huska, J. A., & Keane, T. M. (1993, October). The PTSD Checklist: Reliability, validity, and diagnostic utility. Paper presented at the Annual Meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Weiss, D. S. (2004). The Impact of Event Scale-Revised. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (2nd ed., pp. 168–189). New York: Guilford Press.

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale-Revised. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399–411). New York: Guilford Press.

Woodward, S. H., Murburg, M. M., & Bliwise, D. L. (2000). PTSD-related hyperarousal assessed during sleep. *Physiological Behavior, 70,* 197–203.

Received February 25, 2009
Revision received June 1, 2009
Accepted July 8, 2009 ■